

LSI DOCKET NO. 03-2597

APPLICATION FOR LETTERS PATENT OF THE UNITED STATES

CERTIFICATE OF MAILING BY "EXPRESS MAIL"

"EXPRESS MAIL" Mailing Label Number EV333420216US

Date of Deposit: March 24, 2004

I HEREBY CERTIFY THAT THIS CORRESPONDENCE, **CONSISTING OF 15 PAGES OF SPECIFICATION AND 5 PAGES OF DRAWINGS**, IS BEING DEPOSITED WITH THE UNITED STATES POSTAL SERVICE "EXPRESS MAIL POST OFFICE TO ADDRESSEE" SERVICE UNDER 37 CFR 1.10 ON THE DATE INDICATED ABOVE AND IS ADDRESSED TO: MAIL STOP PATENT APPLICATION, COMMISSIONER FOR PATENTS, P.O. BOX 1450, ALEXANDRIA, VA 22313-1450.

BY: Amy Miller
Amy Miller

SPECIFICATION

To all whom it may concern:

Be It Known, That I, **David Thomas**, a citizen of the United States of America, residing at **12841 NCR 9, Wellington, Colorado 80549**, have invented certain new and useful improvements in "**CRC Data Protection Scheme for Non-Block-Oriented Data**", of which I declare the following to be a full, clear and exact description:

BACKGROUND OF THE INVENTION

1. Technical Field:

The present invention is directed generally toward a method and apparatus for providing
5 hardware-controlled data integrity for data that does not have a fixed block length.

2. Description of the Related Art:

When writing to a device such as a tape, it can be necessary to provide buffering between
a bus and an actual device to smooth out the bursts from the bus into a steady flow onto the tape.
10 This can be handled by a double data rate (DDR) memory device, which allows data to be written
and read from the same memory at different rates. Off-the-shelf DDR memory devices are
available in a variety of sizes that vary both in the number of bits that can be accessed in parallel
and in the number of bits that can be stored. In order to support a high data rate, it is desirable to
have the capability to handle a large number of bits in parallel, e.g., 32 bits.

15 As in all forms of data transmission and storage, whether temporary or permanent, it is
necessary to ensure the integrity of the data during the short time that it resides on the DDR
memory device. This data integrity can be ensured by one of several means.

One exemplary means is by the addition of a single parity bit to a fixed-size set of bits,
with the value of the parity bit determined by the values of the set of bits. To add an odd parity to
20 a string of n bits, a single bit can be added to ensure that the number of ones in the string is odd.
For example, to transmit the four-bit string 1100, one adds a 1 parity bit at the end and transmits
the string 11001. If an error occurs, e.g. the string is received as 11011, the receiving device
determines that the total number of ones is even, so it is clear that an error has occurred.
However, using the single parity bit, it is impossible to determine which bit is in error, so the
25 data must be resent.

Error correction codes (ECCs) carry this one step farther, by providing enough
information that the error can be pinpointed by location and corrected, as long as there are only a
limited number of errors in the group of bits. The downside of an ECC is that it adds a number of
bits to the data stream. For example, using the Hamming code, a common error correction code,

LSI DOCKET NO. 03-2597

three parity bits must be added to four bits of data to ensure that the location of an error can be pinpointed. Luckily, the ratio of data to parity does not remain quite this high, although it is still significant: a four-bit Hamming code will protect eight bits of data; a five-bit code will protect twenty-four bits of data; a six-bit code will protect twenty-eight bits of data, and a 7-bit code will
5 protect thirty-two bits of data. Due to the structure of this code, it can correct a single error or detect several errors, but cannot correct a large number of errors. Other EEC codes will correct n errors and detect $n + m$ errors, with n and m determined by the code formula and number of bits protected.

A cyclical redundancy code (CRC) is another technique, used to protect blocks of data,
10 called frames. For each frame, an n -bit sequence, called a Frame Check Sequence (FCS) is added to the frame. The FCS is calculated such that the new frame, with FCS appended to the original frame, is exactly divisible by a polynomial. When the data is used, the new frame is divided by the same polynomial; if the remainder is not zero, there is an error in the frame.

In order to provide data protection during temporary storage in a DDR memory device, as
15 described above, any of these approaches can be used, if one is willing to deal with the drawbacks. For example, when using a Hamming code with a high data-rate, it becomes necessary to provide either two 32-bit or three 16-bit DDR memory devices in order to pass through both the data and the added code.

In the same situation, a cyclical redundancy code that works on frames would require
20 only one 32-bit DDR memory device. Typically, in a DDR memory device, the application uses a single frame check sequence (FCS) for a large block, such as an entire SCSI command transfer. However, when working with data that is not in fixed lengths, the FCS must be generated and checked by software, a much slower process. With the slower capabilities of software, it is tempting to generate an FCS for as large a block as possible, although this degrades the
25 protection capabilities.

It would be desirable to provide a method of protecting data that is not organized in fixed-length blocks, where the method would provide a high degree of reliability using only one DDR memory device and without resorting to a software-controlled CRC.

LSI DOCKET NO. 03-2597

SUMMARY OF THE INVENTION

The method and device of the invention provide for hardware-controlled data protection that can be used on a device providing buffering between two different protocols, especially where
5 at least one of the protocols does not use known-length blocks. A fixed block size is arbitrarily imposed on the data as it enters the buffering device, in order to calculate a cyclical redundancy code (CRC) for the block. Block sizes are restricted to a value of 2^n , e.g., 2, 4, 8, 16, etc. The device is able to multiplex between ports while sharing the CRC engine between the ports. Intermediate values of the CRC for any given port are temporarily saved in a CRC register file. As a block of
10 data for a given port is completed, a final CRC value for the block is stored in a random access memory (RAM) and the register file entry is cleared. When the data is then output from the device, the CRC for the block is recalculated and checked against the saved value to be sure that they match.

LSI DOCKET NO. 03-2597

BRIEF DESCRIPTION OF THE DRAWINGS

The novel features believed characteristic of the invention are set forth in the appended claims. The invention itself however, as well as a preferred mode of use, further objects and advantages thereof, will best be understood by reference to the following detailed description of an illustrative embodiment when read in conjunction with the accompanying drawings, wherein:

Figure 1 is a block diagram of a computer system on which an embodiment of the invention can be practiced.

Figure 2 is a block diagram of the elements of the device that buffers the data transfer.

Figures 3A-3B together form a flowchart of the actions taken in maintaining the data integrity on the input side of the device.

Figure 3C-3D together form a flowchart of the actions taken in maintaining the data integrity on the output side of the device.

DETAILED DESCRIPTION

An embodiment of the invention will now be explained with reference to the drawings, in which the invention is embodied in a protocol interface chip that is connected to a DDR SDRAM (double data rate synchronous, dynamic random access memory) device. **Figure 1** shows a pictorial representation of a portion of a data processing system **100** in which the present invention may be implemented. In this drawing, processor **102** is connected by bus **104** to memory **106**, disk drives **108**, **110**, and tape drive **114**. Tape drive **114** is connected to the bus through protocol interface device **116**, which is also connected to DDR memory device **112**. Bus **120** connects protocol interface **116** to additional disk drives **122** and **124**. It will be understood that this diagram does not show all the elements of data processing system **100** or all the devices connected to bus **104** and bus **120**, but only enough to clarify the relationship between claimed elements.

Figure 2 is a block diagram of the elements of protocol interface device **116** and DDR memory device **112** of **Figure 1**. In this exemplary embodiment, device **116** has two ports **202R**, **202W** connected to bus **104** and the devices connected to it and two ports **204R**, **204W** connected to bus **120**. These four ports can read and write to devices on their respective bus, e.g., ports **202R** and **202W** are used to read and write to bus **104** while ports **204R** and **204W** are used to read and write to bus **120**. At the same time, port **206** is connected to tape drive **114**. Port **206** can be used as an input port when the tape is read or as an output port when the tape is written. A sixth port **208** is supplied for software data and is used when software control is used rather than full hardware control. A selection device **SEL** connects one of the ports **202R**, **202W**, **204R**, **204W**, **206**, **208** at a time to the CRC engine **210**. As input from the ports is received, the data is stored in blocks of memory that have a length of 2^n . The CRC engine **210** uses the data to calculate a cyclical redundancy code (CRC) value for each block. The CRC engine **210** maintains a running CRC value for each port while the data is being stored; as the CRC engine **210** switches between ports, these intermediate values are stored in the CRC register file **212**. As each block is completed, a final CRC value for the block is written to CRC RAM **214**, where it is maintained until it is overwritten, and the corresponding entry in the register file **212** is reset.

LSI DOCKET NO. 03-2597

Information in the CRC RAM **214** is stored in a mapped relationship to the corresponding data that is stored in a data memory. When the data is being read from the DDR SDRAM, the CRC engine **210** will be given the data values stored in memory so that it can again calculate a CRC value. This value is fed to the comparator **216**, where the newly calculated CRC value is
 5 compared to the CRC value that was previously stored in the CRC RAM. If the two values match, the data is correct; if not, an error has occurred and corrective action is initiated. Results from the comparator **216**, as well as data stored in the CRC RAM **214**, are sent to the microprocessor **218**, where they can be written from appropriate ports.

The actions taking place in the interface protocol chip will now be explained with
 10 reference to **Figures 3A-3D**. As these actions are discussed, it is helpful to keep in mind the following limitations: 1) stored data must begin and end on a block boundary and 2) any padding can only be done at the beginning and/or end of a transfer. Additionally, for the purposes of this discussion, data is stored in the memory in blocks 1024 bytes (2^{10} or 1K) long. As mentioned above, the CRC engine is shared by both the input and output side of the device and may support
 15 more than one port at a time. The selector switch **SEL** is used to determine which port is connected to the CRC engine at any given time. **Figure 3A**, which is continued on **Figure 3B**, shows the flowchart for protecting input data. This flow begins when the selector sends input from a port, here referred to as **port_p**, to the CRC engine. When the device receives input from a given protocol, the data will generally have an embedded protection code according to the
 20 protocol from which it is coming, e.g. the bus. Rather than passing this protection code through the memory device, which can be expensive in terms of memory space, bandwidth and logic complexity, the protection code is first checked to ensure that the data is correct, then the protection code is discarded. When the CRC engine is entered, it is necessary to determine if this is a new transmission for **port_p** or the resumption of a transmission that was not previously
 25 completed (step 302). If it is a new transmission, blocks for storage will be allocated and values will be set for the variables $\mathcal{L}_{transfer}$, $\mathcal{L}_{storage}$, **b**, **addr(b)**, **B**, **s**, **i**, **x**, and **y**, as explained below (step 306). The transfer length, $\mathcal{L}_{transfer}$, is determined by the source of the data, in this example, either a bus or a tape, and will be received as input. For hardware to manage the protection, the stored data must begin and end on a block boundary. Since the blocks are 1K long in this example,

LSI DOCKET NO. 03-2597

$\mathcal{L}_{\text{storage}}$ will be set to the lowest multiple of 1K that is equal to or greater than $\mathcal{L}_{\text{transfer}}$. If $\mathcal{L}_{\text{transfer}}$ is not equal to $\mathcal{L}_{\text{storage}}$ (as happens most of the time), the additional space must be padded with known values, such as zeros. If, for example, $\mathcal{L}_{\text{transfer}} = 6.5\text{K}$, seven blocks will be allocated and $\mathcal{L}_{\text{storage}} = 7\text{K}$. Variable **B** is set to the number of blocks used, in this case seven (7), while **b** is the
 5 number of the current block being used and is set to one (1); **addr(b)** refers to the beginning address of the current block and will be initialized to the address of the first block allocated. Variable **s** is an offset into the current block, while variable **i** is the offset into the input buffer; both are initialized to zero (0). Finally, variables **x** and **y** define the amount of padding needed at the beginning and end of the storage respectively. When handled by hardware, **x** is generally zero
 10 (0), although this is not necessary to the method; **y** is always equal to $((\mathcal{L}_{\text{storage}} - \mathcal{L}_{\text{transfer}}) - \mathbf{x})$. In the example, if **x** = 0, then **y** = 512. Because of the limitation that padding can only be done at the beginning and end of the transmission, it is only necessary to check the value of **x** if **b** = 1 and to check **y** only if **b** = **B**. The hardware is enabled to set the values of all these variables.

If this is not a new transmission, but the resumption of a previous transmission that was
 15 not completed, the values for the variables, **b**, **B**, **s**, **i**, **x**, **y**, $\mathcal{L}_{\text{transfer}}$, and $\mathcal{L}_{\text{storage}}$ will have been previously stored and are retrieved (step 308).

In order that the CRC engine can be interrupted, it maintains its calculations-in-progress in a register file, saved according to the port being used. The CRC engine retrieves this value, **CRC(port_p)**, at the beginning of the calculations (step 310); for a new transmission, this value
 20 will be a null value.

The CRC engine first checks if the value of (**b** = 1 and **s** is less than **x**), i.e., if there is any padding necessary at the beginning of the stored data (step 312). If the answer is “yes”, a byte of padding, which in this example is zero, is stored in the data memory at (**addr(b)** + **s**) (step 314). If the answer is “no”, the engine checks further if (**b** = 1 and **s** \geq **y**), i.e., if padding is necessary at
 25 the end of the stored data (step 316). An answer of “yes” here will also cause the engine to store padding in the data memory at (**addr(b)** + **s**) (step 314). If, however, the answer to the questions of both step 312 and step 316 are ‘no’, input byte (**i**) is stored at (**addr(b)** + **s**) of the memory and **i** is incremented by 1 (step 318). Regardless of what was stored at offset **s**, the CRC engine calculates a new CRC value, incorporating the latest stored value into the previous calculations,

LSI DOCKET NO. 03-2597

then, the value of s is incremented by 1 (step 320).

Once the CRC value has been updated, it is necessary to check whether the stored information has reached a block boundary (i.e. whether $s = 1024$ in this example) (step 322), as this will determine whether the CRC value is a temporary value or a final value. If the end-of-block has been reached (“yes”), the CRC calculation is complete and the CRC value will be written to CRC memory (step 324). This memory is located on memory device 112, but can be separate from the data memory. The CRC value is stored at a mapped location with respect to the block it protects, facilitating hardware handling of the coding. Once this is done, the register file for **port_p** is reset to a null value and the CRC engine is ready to proceed. Since the current block **b** has just been updated, a check to see if $b > B$ (the total number of blocks) is next made (step 328). If the current block is greater than the total number of blocks, this input is complete and the flowchart is exited. If not, the flowchart goes to step 332, which will be discussed below. Returning to step 322, if a block boundary was not found, the new intermediate value of CRC (**port_p**) is written to the register file so that it will not be lost if the selector shifts to another port (step 326) and this branch also proceeds to step 332.

If the selection switch chooses to switch ports, it can do so at this point (step 332). If the selector switch is changing ports, the values of **b**, **B**, **addr(b)**, **i**, **s**, **x**, **y**, $L_{storage}$, and $L_{transfer}$ are saved for **port_p** (step 330) and the flowchart is exited. Otherwise, the flow returns to step 312 to continue processing the present transmission.

Next, **Figure 3C**, which is continued on **Figure 3D**, shows the flowchart for checking output data read from the DDR SDRAM memory chip. As in the previous flow, all the ports share a single CRC engine, so information for each port must be carefully saved so that the values are accessible when processing for that port resumes. The flow begins when processing for **port_p** is initiated. When entered, the CRC engine first checks whether this is a new transmission for **port_p** or this is a continuation of a previous transmission (step 354). If this is a new transaction, the values of **b**, **B**, **addr(b)**, **i**, **s**, **x**, **y**, $L_{storage}$, and $L_{transfer}$ are set (step 356). If this is a continuation of a previous transmission, these values will be retrieved from where they were previously stored (step 358). The value of **CRC (port_p)** is retrieved from the register file (step 360). If ($b = 1$ and $s < x$) (step 362) or if ($b = B$ and s is equal or greater than y) (step 364),

LSI DOCKET NO. 03-2597

then the current byte was previously padded and the byte will not be moved to the output buffer. If neither condition is true, the byte at (**addr(b)** + **s**) is moved to the output buffer (**i**) and **i** is incremented (step 366). Regardless of whether the byte was moved to the output buffer or not, it is used to calculate the CRC value, the new value is saved to **CRC(port_p)**, and **s** is incremented by 1 (step 368).

In order to check that the data being output is correct, the next step is to determine whether the end of a block boundary has been reached, e.g., $s = 2^n$ (step 370). If it has not, the next few steps will be skipped and the flow proceeds to step 378. If the end of a block has been reached, the next step is to check to see if the newly calculated CRC value matches the value stored in the CRC random access memory when the data was stored (step 372). If the two values match, the data thus far is correct and can be sent. The stored value of **CRC(port_p)** in the register file is reset, **b** is incremented by 1 and **s** is set to 0 (step 376) to prepare for processing the next block, otherwise an error recovery process is initiated according to protocol (step 374). The CRC value that was stored in the mapped location will remain intact until the space is next needed. Next, the engine checks for the completion of the current transmission, e.g., if **b** > **B** (step 378). If it is, the flowchart is exited. If further data remains in this transmission, it is necessary to determine if the selector will be switching to process another port (step 380). If it is, the values of **b**, **B**, **addr(b)**, **i**, **s**, **x**, **y**, **L_{storage}**, and **L_{transfer}** are saved for this port (step 382) and the flow ends. If not, processing will return to step 362 and continue.

The foregoing description has shown a device and method for maintaining data protection as data is passed between two protocols. The advantages of this method and device are that the protection can be entirely hardware based, even when the data transfers are of arbitrary lengths. This embodiment minimizes the cost of the buffer portion of the system and uses a single CRC engine for multiple data flows.

The description of the preferred embodiment of the present invention has been presented for purposes of illustration and description, but is not intended to be exhaustive or limited to the invention in the form disclosed. Many modifications and variations will be apparent to those of ordinary skill in the art. For example, the number of ports on the device can vary between embodiments; e.g., in alternate embodiments, the CRC engine can support multiple ports on both

LSI DOCKET NO. 03-2597

sides of the device. Additionally, the busses shown can be a fabric of connections, rather than single busses. The exemplary embodiment was shown with padding equal to zero. In other embodiments, other padding can be used. For example, the padding may consist of data read during tape read-ahead. In another alternative, the CRC register file 212 can be part of the CRC
5 RAM 214 if data path bandwidth is sufficient. Additionally, although the invention was described using a DDR SDRAM placed between a bus and a tape drive, it is not limited to this embodiment. In alternate embodiments, the invention can be practiced in a device such as bridge that maps different protocols with different block sizes. Alternatively, the invention can be practiced in a buffer used to translate RapidIO to Fibre Channel. The embodiment was chosen
10 and described in order to best explain the principles of the invention the practical application to enable others of ordinary skill in the art to understand the invention for various embodiments with various modifications as are suited to the particular use contemplated.